IDA PAPER P-2609

RELATING FLYING HOURS TO AIRCREW PERFORMANCE: EVIDENCE FOR ATTACK AND TRANSPORT MISSIONS

Colin P. Hammon
Stanley A. Horowitz, *Project Leader*

June 1992

Prepared for
Office of the Assistant Secretary of Defense
(Force Management and Personnel)

INSTITUTE FOR DEFENSE ANALYSES
1801 N. Beauregard Street, Alexandria, Virginia 22311-1772

IDA Log No. HQ 91-39384

DEFINITIONS

IDA publishes the following documents to report the results of its work.

Reports

Reports are the most authoritative and most carefully considered products IDA publishes. They normally embody results of major projects which (a) have a direct bearing on decisions affecting major programs, (b) address issues of significant concern to the Executive Branch, the Congress and/or the public, or (c) address issues that have significant economic implications. IDA Reports are reviewed by outside panels of experts to ensure their high quality and relevance to the problems studied, and they are released by the President of IDA.

Group Reports

Group Reports record the findings and results of IDA established working groups and panels composed of senior individuals addressing major issues which otherwise would be the subject of an IDA Report. IDA Group Reports are reviewed by the senior individuals responsible for the project and others as selected by IDA to ensure their high quality and relevance to the problems studied, and are released by the President of IDA.

Papers

Papers, also authoritative and carefully considered products of IDA, address studies that are narrower in scope than those covered in Reports. IDA Papers are reviewed to ensure that they meet the high standards expected of refereed papers in professional journals or formal Agency reports.

Documents

IDA Documents are used for the convenience of the sponsors or the analysts (a) to record substantive work done in quick reaction studies, (b) to record the proceedings of conferences and meetings, (c) to make available preliminary and tentative results of analyses, (d) to record data developed in the course of an investigation, or (e) to forward information that is essentially unanalyzed and unevaluated. The review of IDA Documents is suited to their content and intended use.

The work reported in this document was conducted under contract MDA 903 89 C 0003 for the Department of Defense. The publication of this IDA document does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official position of that Agency.

UNCLASSIFIED

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
June 1992

3. REPORT TYPE AND DATES COVERED
Final Report, Sep 1991 – Jun 1992

4. TITLE AND SUBTITLE
Relating Flying Hours to Aircrew Performance: Evidence for Attack and Transport Missions

5. FUNDING NUMBERS
C-MDA-903-89C-0003
T-L7-516

6. AUTHOR(S)
Colin P. Hammon and Stanley A. Horowitz

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Defense Analyses
1801 N. Beauregard Street
Alexandria, VA 22311-1772

8. PERFORMING ORGANIZATION REPORT NUMBER
IDA-P-2609

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
OASD(FM&P)
Room 3E763, The Pentagon
Washington, D.C. 20301

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12A. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12B. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This paper describes the development of quantitative relationships between the flight experience of military flying personnel and how well they perform important aspects of their mission. This research responds in part to concerns expressed by the General Accounting Office and Congress about the impact of reductions in the services' flying-hour programs. Research confirms the hypothesis that experience enhances proficiency in two ways: through the short-run honing of skills and through the long-run development of mastery. Measures of experience include flights and flying hours in the previous 7 to 60 days and the number of career flying hours. Econometric models and estimates of the strength of the links between these experience measures and two measures of performance are developed. The performance measures are bombing and tactical airdrop accuracy. In addition, the tradeoff between flying and flight simulator hours is examined for both Marine Corps and C-130 crewmembers. We conclude that although short-run measures are significant, the most important determinant of proficiency is total flying hours. This means that in an emergency it would be difficult to correct long-term deficiencies in accumulated flying hours.

14. SUBJECT TERMS
Military Personnel, Pilots, Flying Hours, Flight Simulators, Air Drop Operations, Tactical Bombing, Accuracy

15. NUMBER OF PAGES
39

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

UNCLASSIFIED

IDA PAPER P-2609

RELATING FLYING HOURS TO AIRCREW PERFORMANCE: EVIDENCE FOR ATTACK AND TRANSPORT MISSIONS

Colin P. Hammon

Stanley A. Horowitz, *Project Leader*

June 1992

INSTITUTE FOR DEFENSE ANALYSES

Contract MDA 903 89 C 0003

Task T-L7-516

PREFACE

This paper was prepared by the Institute for Defense Analyses (IDA) for the Office of the Assistant Secretary of Defense (Force Management and Personnel) under contract MDA 903 89C 0003, Task Order T-L7-516, issued 17 June 1987, and amendments. The objective of the task was to facilitate development of quantitative relationships between the capability of aviation units to perform their assigned missions and the levels of training resources (flying hours and simulator time) available to them.

This work was reviewed within IDA by Matthew S. Goldberg, Philip M. Lurie, and Jesse Orlansky. Colin P. Hammon, one of the authors of this paper, is an IDA consultant.

CONTENTS

FIGURES

TABLES

I. INTRODUCTION

Over $10 billion a year is spent on the flying-hour program of all the military services. The flying-hour program is the key element in keeping pilots trained, but there is considerable doubt about the validity of the way in which the flying-hour program is developed and about the value of maintaining flying hours at the current or other levels. Flying hours are an attractive target for congressional budget reductions because money can be saved immediately, as opposed to the longer spend-out periods for acquisition programs. In an era of lower international tensions, there will be a tendency to reduce flying-hour programs, perhaps dramatically. In this environment, it is important to have a better idea of just what the implications of variations in flying hours are for the performance of U.S. pilots, and for our ability to effectively mobilize aviation assets.

This paper develops quantitative relationships between how much aircrews fly and how well they perform important aspects of their missions. The research described here responds in part to concerns expressed by the General Accounting Office and to congressional skepticism about the impact of reductions in the services' flying-hour programs [1 and 2].

An earlier paper described the small body of existing literature that has developed such relationships [3]. The following overall conclusions were reached:

- Quantitative relationships that support the proposition that more flying results in measurably better performance have been developed for both the Air Force and the Navy.
- Additional flying appears to improve aircrew performance in two ways. In the short run, it hones skills and prevents their deterioration. In the long run, it permits the attainment of a higher level of mastery that is reflected in better performance. None of the existing analyses that were reviewed fully captures both of these effects.
- Data exist to develop additional links between flying-hour activity and measures of operational performance for a wide range of aircraft. Additional research to build such links should be done.

The results of new statistical analyses that examined both the long-range and short-range effects of flying hours on performance were reported in Reference [4]. Three

empirical investigations were reported. The first examined the quality of landings aboard aircraft carriers for F-14 and A-7 aircraft. The second focused on the accuracy with which Marine Corps aviators dropped bombs from AV-8, F/A-18, and F-4S aircraft. The third drew on the performance of F-14 fighters during opposed air-combat-maneuvering exercises on an instrumented range.

The findings confirmed the existence of both short- and long-run positive effects of flying hours on aircrew performance. Additionally, the relative magnitudes of these effects were estimated. Table 1 summarizes the impact on performance of a 10% decrease in flying hours from the current level, as reported in [4].

Table 1. Impact on Performance of 10% Cuts in Flying-Hour Variables (Previous Analyses)

Performance Measure	Career Flying	Recent Flying	Total
Unsatisfactory Landings	6.9%	2.6%	9.5%
Bombing Miss Distance	1.5%	1.0%	2.5%
Air-to-Air Combat			
Probability Red Kills Blue	6.3%	2.9%	9.2%
Probability Blue Kills Red	-2.6%	-2.2%	-4.8%

Source: Reference [4].
Notes: Red are aggressor aircraft; Blue are friendly aircraft.

This study reports the results of additional statistical analyses of both the long- and short-term effects of flying hours on performance. Two empirical investigations are reported. The first examines the accuracy with which Marine Corps aviators dropped bombs from AV-8, F/A-18, and F-4S aircraft. This analysis re-examines a research area reported in [4] using an expanded database. The second addresses the performance of C-130 aircrews flying tactical airlift training missions. These analyses extend the previous work in an important way. In addition to actual career and recent flying hours, both career and recent simulator hours were available for analysis. Additionally, the C-130 is a multiperson aircraft, and it was possible to isolate the flying-hour effects on total mission performance of different crewmembers.

Our findings confirm the existence of both short- and long-term positive effects of flying hours on aircrew performance. We were also able to estimate the effect on performance of Marine Corps pilot and C-130 co-pilot simulator hours.

The analyses and results of these investigations are discussed in Sections II and III. Conclusions and recommendations are stated in Section IV.

II. ANALYSIS OF MARINE CORPS BOMBING

The analysis reported here represents an enhancement of work reported in Reference [4]. Because we were able to assemble a larger and more detailed database, the results reported here are more detailed and robust than those of the previous analysis. Of particular importance, we were able to isolate the effect on performance of flight simulator hours.

Marine Corps Headquarters maintains a database that includes the information on flying-hour history and, in some cases, performance needed to examine the existence of statistical relationships between flying hours and aircrew performance. This information is part of the Naval Flight Record Subsystem (NAVFLIRS). Pilots fill out flight activity reports after every flight. When this activity includes bombing practice on instrumented ranges, the accuracy with which bombs were dropped may be included in the report. The master NAVFLIRS database maintained at Marine Corps Headquarters also includes information on the career flying and flight simulator history of each pilot.[1]

A. DATA

Performance data were obtained for 1,003 bombing exercises flown by AV-8 and F/A-18 fleet and fleet replacement training squadrons. These were supplemented with performance data for 738 bombing exercises recorded directly on the NAVFLIRS reports. The database is derived from bombing runs of three kinds of aircraft, AV-8, F/A-18, and F-4.

Both manual bomb deliveries and deliveries in which a computer fire-control system was employed are included, and every bombing exercise is identified as either manual or computed. Computed deliveries are further broken down into automatic and Continuously Computed Impact Point (CCIP), two different modes within the capability of

---

[1] The Navy also uses NAVFLIRS, but when practice bombing is performed, Navy pilots do not report bombing accuracy results to the database. Marine pilots did report bombing accuracy for a period of time when the system was first implemented. In addition, one squadron recorded bombing accuracy on the NAVFLIRS report for this study. NAVFLIRS has been in use since January 1987.

the AV-8 and F/A-18 fire-control systems.[2] Each exercise (observation) consists of a series of from 1 to 18 bombing runs by a single pilot on a single sortie where the same type of delivery (dive angle, delivery mode—automatic, CCIP, or manual—ordnance, etc.) is performed. More than one exercise may be flown on a single sortie. Experience data were taken, for all 1,741 observations, from the Headquarters database. Table 2 summarizes the Marine Corps data.

Table 2. Data Used in the Analysis of Marine Corps Bombing

Variable	Minimum	Mean	Maximum
Median Miss Distance (feet)	0	102	1,583
Number of Runs	1	4.6	18
Career Flying Hours	255	1,274	5,875
Career Flight Simulator Hours	0	79	256
Flights in the Previous 7 Days	0	3.2	10
Flying Hours in the Previous 7 Days	0	4.1	22
Simulator Hours in the Previous 7 Days	0	0.28	5.1
AV-8 Automatic Deliveries[a]	0	0.20	1
F/A-18 Automatic Deliveries	0	0.25	1
Total Automatic Deliveries	0	0.44	1
AV-8 CCIP Deliveries	0	0.23	1
F/A-18 CCIP Deliveries	0	0.11	1
Total CCIP Deliveries	0	0.34	1
AV-8 Manual Deliveries	0	0.11	1
F/A-18 Manual Deliveries	0	0.05	1
F-4 Manual Deliveries	0	0.07	1
Total Manual Deliveries	0	0.22	1
AV-8 Flight	0	0.53	1
F/A-18 Flight	0	0.40	1
F-4 Flight	0	0.07	1
MK-76 Practice Bombs	0	0.77	1
Loft Delivery	0	0.03	1
Fleet Replacement Pilot	0	0.60	1

[a] The last 16 variables in the table are dichotomous. They either take the value one, showing that an observation has the indicated property, or the value zero, showing that it does not. The mean is the fraction of the observations that have the indicated property.

2 In the automatic mode, a target designation symbol is placed over the target on the pilot's head-up display (HUD). The aircraft is then flown to the bomb release point where the bombs are automatically released. In the CCIP mode, the impact point is continuously computed and displayed on the HUD. The pilot steers the aircraft so that the display representation of the impact point is superimposed over the target, and releases the bombs. Both the automatic and CCIP modes are more complicated than this explanation implies, and the pilot must also monitor many parameters, such as dive angle, g-loading, altitude, attitude, fuzing, etc.

Three-quarters of the bombing observations represent computed deliveries. This reflects the fact that a large majority of the AV-8 and F/A-18 deliveries were computed. The F-4 does not have a computer delivery system. Of the 1,361 computed observations, 774 were automatic and 587 were CCIP. Approximately one-half of the automatic and one-third of the CCIP observations are for F/A-18 flights.

MK-76 practice bombs were used on three-quarters of the missions and actual bombs, or inert shapes—sand- or water-filled replicas—were dropped on the remaining flights. The loft maneuver, which is generally less accurate than other deliveries, was used on 3% of the observations. Sixty percent of the missions were flown by Fleet Replacement Pilots (FRPs). These are pilots who are receiving refresher training prior to reporting to a fleet squadron. They have either recently completed undergraduate flight training, have been in a non-flying billet, or are changing to a different aircraft type.

In addition to actual flying-hour data, the database included flight simulator hours, both recent and career, for each pilot we observed.

The median miss distance of a series of bombing runs was chosen as the measure of performance for two reasons. First, we did not have individual drop scores for the 738 observations recorded directly on the NAVFLIRS flight reports. Only the delivery type, ordnance, number of runs, and the median miss distance are recorded on the NAVFLIRS form. In order to use individual scores, we would have had to eliminate these 738 observations, which included all of the F-4 and most of the fleet pilot observations. Second, the median is the circular error probable (CEP);[3] hence, it better represents the probability of kill for a multibomb delivery. This is the reason given by the Marine Corps for recording the median rather than the mean miss distance on the NAVFLIRS flight report.

B. ANALYSIS

The central hypotheses of this analysis are that pilots with more career flying and flight simulator experience drop their bombs more accurately and that greater recent flying and simulator experience is associated with more accurate bombing. Subsidiary hypotheses are that computed bomb deliveries are more accurate than manual ones, and that experience—both short-term and long-term—plays a smaller role in determining the accuracy of computed deliveries than of manual deliveries. One of the effects of computer

---

[3] The median (Me) of a continuous random variable X is defined as $P[X \leq Me] = 1/2$

fire-control systems should be to bring pilots with less experience to a high state of proficiency in a shorter time. Computer delivery systems are meant to increase accuracy.

It is also expected that bombing accuracy will differ according to the type of aircraft. The newest aircraft, the F/A-18, is expected to be the most accurate, and the oldest, the F-4, the least accurate. These hypotheses led to the formulation of the following equation:[4]

$$\begin{aligned} LnCE = {} & b_0 + b_1 \times LnH_c \times M + b_2 \times LnH_c \times A + b_3 \times LnH_c \times C + b_4 \times LnH_{cs} \\ & \times (A + C) + b_5 \times LnF_7 \times M + b_6 \times LnH_{7s} \times (1 - R) + b_7 \times A \times F18 \\ & + b_8 \times C \times F18 + b_9 \times M \times F18 + b_{10} \times A \times AV8 + b_{11} \times C \times AV8 \\ & + b_{12} \times M \times AV8 + b_{13} \times R + b_{14} \times B_{76} + b_{15} \times L \end{aligned} \qquad (1)$$

where:

- Ln = the natural log
- CE = miss distance (circular error), the distance in feet by which the bomb misses the target (CE is the median for a series of bombing runs. The number of runs varied between 1 and 18.)
- H_c = career flying hours
- H_{cs} = career flight simulator hours
- F_7 = flights in the previous 7 days[5,6]
- H_{7s} = flight simulator hours in the previous 7 days
- A = a dummy variable taking the value 1 for automatic deliveries and zero otherwise
- C = a dummy variable taking the value 1 for CCIP deliveries and 0 otherwise
- M = a dummy variable taking the value 1 for a manual delivery and 0 otherwise
- AV8 = a dummy variable taking the value 1 for an AV-8 flight and 0 otherwise
- F18 = a dummy variable taking the value 1 for an F/A-18 flight and 0 otherwise

4 Several variants of Equation (1) were estimated using the full database as well as various subsets (i.e., manual only, computed only, computed F/A-18 or AV-8, etc.) The results of these analyses were quite robust and did not depend on the particular sample. These excursions led us to Equation (1).

5 In a previous analysis of carrier landing grades, the short-term experience variable was flying hours in the previous month [4]. The 7-day variable was more successful in Equation (1) than the previous month variable. There is no intrinsic reason why recently honed skills ought to depreciate at the same rate for landing and bombing. It is, however, interesting that the bombing analysis was more sensitive to the choice of a short-term experience variable than the landing grade analysis.

6 We estimated the equation using both flights and hours in the previous 7 days. The results were essentially the same.

R = a dummy variable taking the value of 1 for FRPs and 0 for fleet pilots

B_{76} = a dummy variable taking the value 1 for MK-76 practice bombs and 0 otherwise

L = a dummy variable taking the value 1 for loft deliveries and 0 otherwise

b_0 through b_{15} are coefficients to be estimated.

The coefficients b_1, b_2, and b_3 measure the effect of additional career flying hours on bombing performance for manual, automatic, and CCIP observations, respectively. All three coefficients are expected to be negative; better performance is reflected in smaller miss distances. The coefficient b_1 is expected to be greater in absolute value than either b_2 or b_3 since we conjecture that the long-term effect is greater for manual than for computed deliveries.

Similarly, b_4, the elasticity of bombing accuracy with respect to career flight simulator hours for computed approaches, should be negative. Although we conjecture a substitution effect between flying hours and flight simulator hours, we believe that simulator hours are much more closely related to computed deliveries than to manual deliveries. Most of a pilot's career simulator time is spent in relatively unsophisticated trainers where the emphasis is on learning procedures rather than actual flying skills. Proficiency in procedures associated with the fire-control system is a key to computed bombing. An equation with separate coefficients for computed and manual deliveries was estimated, and the coefficient of career simulator hours was not significant for manual approaches—the t-statistic equalled .2.

We also conjecture that b_5, the coefficient of flights in the previous 7 days, should be negative for manual approaches. The equation with separate coefficients for computed and manual deliveries was estimated, and short-term experience has no observable effect for computed deliveries. The coefficients were statistically insignificant—t-statistics equalled .57 and -.2 for automatic and CCIP deliveries, respectively.

An examination of the data revealed little variation in the value of simulator hours in the previous 7 days for FRPs. This is consistent with the fact that all FRPs go through essentially the same simulator syllabus, whereas recent simulator hours vary a great deal for fleet pilots. We therefore included only recent simulator time for fleet pilots in the estimating equation. The coefficient of this variable, b_6, is expected to be negative.

We expect b_7, b_8, b_{10}, and b_{11} to be negative, picking up the role of computed bombing systems in improving accuracy. We also expect b_7 and b_8 to be greater in absolute value than b_{10} and b_{11}, respectively, because the F/A-18 has a newer, more

sophisticated fire-control system than the AV-8. Although there was not a consensus among the pilots interviewed, many experienced pilots indicated that CCIP deliveries are easier for beginning pilots to execute satisfactorily. We therefore expect greater accuracy for CCIP deliveries than for automatic deliveries. We cannot predict the signs for b_9 and b_{12}. These coefficients measure the difference in accuracy for manual F/A-18 and AV-8 deliveries relative to the F-4. Since the F-4 flies only manual approaches, we might expect this aircraft—the only one of the three that is not explicitly identified by a variable in the equation—to be no less accurate for manual approaches than the AV-8 or the F/A-18.

R indicates a series of bombing runs flown by a Fleet Replacement Pilot. Because FRPs are undergoing refresher training for one reason or another, we expect bomb deliveries by them to be less accurate than deliveries by fleet pilots.

B_{76} and L are control variables. We expect accuracy to be different for MK-76 practice bombs than for shapes or actual bombs, since the wind affects each differently. The loft maneuver is used for simulated nuclear deliveries, where the escape path is more important than pinpoint accuracy. We expect the coefficient b_{15} to be positive.

We initially believed that accuracy might vary with the number of runs since we conjecture that learning takes place during the mission. The number of runs for each observation was included in our preliminary estimates to control for this effect. However, this variable did not contribute significantly to the explanatory power of the equation.

Table 3 includes our expectations regarding the signs of the explanatory variables.

C. RESULTS

Equation (1) was estimated using weighted least squares with the number of runs for each observation as the weight.[7] The equation was estimated in log-log form based on the observed form of the raw data and our *a priori* beliefs about learning. We expect learning to exhibit diminishing returns to the experience variables. The estimated coefficients are reported in Table 3. The results shown in Table 3 are consistent with all of the primary hypotheses discussed above. The results do not support our conjecture of greater accuracy of the F/A-18 relative to the AV-8.

---

7 The estimated equation explains about one-third of the variation in bombing accuracy. This means that the equation cannot very precisely predict where an individual bomb will fall based on the explanatory variables in Equation (1). The goal of this paper is not, however, to predict the location of individual bomb deliveries; rather, it is to estimate the effect of flying hours on the average accuracy of a large number of deliveries. The statistical significance of the coefficients of the independent variables indicates that it is adequate to this task.

Table 3. Determinants of Bombing Accuracy for Marine Corps Aircraft (Median Miss Distance in Feet)

	Dependent Variable: LnCE			
Independent Variable	Coefficient	Expected Sign	Value of Coefficient	t-statistic
Constant			5.00	13.28***
Career Flying Hours for Manual Deliveries ($LnH_c \times M$)	b_1	–	-0.1174	2.86***
Career Flying Hours for Automatic Deliveries ($LnH_c \times A$)	b_2	–	-0.1086	3.46***
Career Flying Hours for CCIP Deliveries ($LnH_c \times C$)	b_3	–	-0.0718	1.86*
Career Flight Simulator Hours for Computed Approaches ($LnH_{cs} \times (A + C)$)	b_4	–	-0.0243	2.96***
Flights in the Previous 7 Days for Manual Deliveries ($LnF_7 \times M$)	b_5	–	-0.0610	2.31**
Simulator Hours in the Previous 7 Days for Fleet Pilots ($LnH_{7s} \times (1 - R)$)	b_6	–	-0.1895	1.85*
F/A-18 Automatic Delivery (F18 × A)	b_7	–	-0.7461	1.95**
F/A-18 CCIP Delivery (F18 × C)	b_8	–	-1.3463	3.25***
F/A-18 Manual Delivery (F18 × M)	b_9	—[a]	-0.1492	1.50
AV-8 Automatic Delivery (AV8 × A)	b_{10}	–	-.9946	2.66***
AV-8 CCIP Delivery (AV8 × C)	b_{11}	–	-1.4655	3.69***
AV-8 Manual Delivery (AV8 × M)	b_{12}	—[a]	-0.3434	4.00***
Fleet Replacement Pilot (R)	b_{13}	+	0.8351	3.57***
MK-76 Practice Bomb (B_{76})	b_{14}	—[a]	0.1891	4.54***
Loft Delivery (L)	b_{15}	+	1.4423	14.41***
Number of Observations = 1,742				
Adjusted R^2 = .32				

[a] We have no hypothesis concerning the signs of these coefficients.
* Significant at the .1 level.
** Significant at the .05 level.
*** Significant at the .01 level.

The coefficients of both the long-term flying-hour and simulator-hour variables are statistically significant in the expected direction. Additionally, the long-term flying-hour effect is significant but of a lower magnitude for computed bombing. For manual bombing, additional flights in the previous 7 days improve performance. For fleet pilots, miss distances also decrease with an increase in simulator hours in the previous 7 days.

As expected, automatic deliveries are more accurate than manual deliveries, and CCIP drops are more accurate than automatic ones for the same type of aircraft. Surprisingly, both AV-8 automatic and CCIP deliveries are more accurate than the

comparable F/A-18 deliveries.[8] F/A-18 manual deliveries are also more accurate than for the F-4. However, this coefficient is statistically significant only at the .13 level.

Table 4 shows the partial effects and elasticities at the mean of the experience variables. The partial effect is the change in the circular error (in feet) associated with a unit change in the independent variable. The elasticity of CE with respect to a given independent variable is the percentage change in CE associated with a 1% change in the independent variable. For the variables that appear in the estimated equation (Table 3), the coefficient is equal to the elasticity. The elasticities shown in Table 4 are calculated as weighted averages of the coefficients of the appropriate independent variables included in Table 3.

Table 4. Partial Effects and Elasticities of Circular Error With Respect to Experience Variables for Marine Corps Bombing

Independent Variable	Partial Effect	Elasticity
Career Flying Hours (H_C)	-0.0079	-0.0982
Career Flight Simulator Hours (H_{CS})	-0.0245	-0.0190
Flights in the Previous 7 Days (F_7)	-0.4263	-0.0134
Simulator Hours in the Previous 7 Days	-28.1533	-0.0765

Some implications of the estimated results shown in Table 3 are presented graphically in Figures 1 through 5.[9,10] Figures 1 and 2 depict the effect of changing career flying hours for fleet pilots and FRPs, respectively.[11] Figure 3 shows the effect of varying career flight simulator hours for computed deliveries. Figure 4 shows the effect of changing flights in the previous 7 days for manual deliveries. Figure 5 shows the effect of changing simulator hours in the previous 7 days for fleet pilots. In all cases, independent variables other than the one being examined are held at their mean values.

---

8 The hypothesis that the coefficients are equal is rejected at the .07 to.0001 level, using sequential F-tests.

9 In order to graph the estimated equation, we made the transformation:

$CE = \exp\{LnCE\} = \exp\{b_0\} \prod_{i=1,6} x_i^{b_i} \exp\{\sum_{i=7,15} b_i x_i\}$.

However, this is a biased estimate of CE because the mean of the transformed residual, which is non-zero, is omitted. In estimating LnCE, we assume that Lnu_i is distributed normal with mean zero and standard deviation s. This means that u_i is distributed lognormal with mean exp $\{s^2/2\}$. We chose the value of s to equate the mean of the predicted equation to the mean of the raw data, yielding the estimate s = 1.04. We therefore include the correction exp{.54}. The transformed equation is therefore:

$CE = \exp\{b_0\} \prod_{i=1,6} x_i^{b_i} \exp\{\sum_{i=7,15} b_i x_i\} \exp\{.54\}$.

10 In Figures 1–4, the first, third, and fifth points on the graphs correspond to the minimum, mean, and maximum values for the independent variable, and the second and fourth points correspond to the midpoints between those values. In cases where the minimum value was zero, a small constant was added to prevent the predicted value of the dependent variable from being infinite. In Figure 5, points were chosen in a somewhat different fashion to display the effect of variation over the range of the independent variable.

11 The minimum, mean, and maximum values of career flying hours are greater for fleet pilots than for FRPs. This occurs because a greater percentage of the FRPs are first tour pilots.

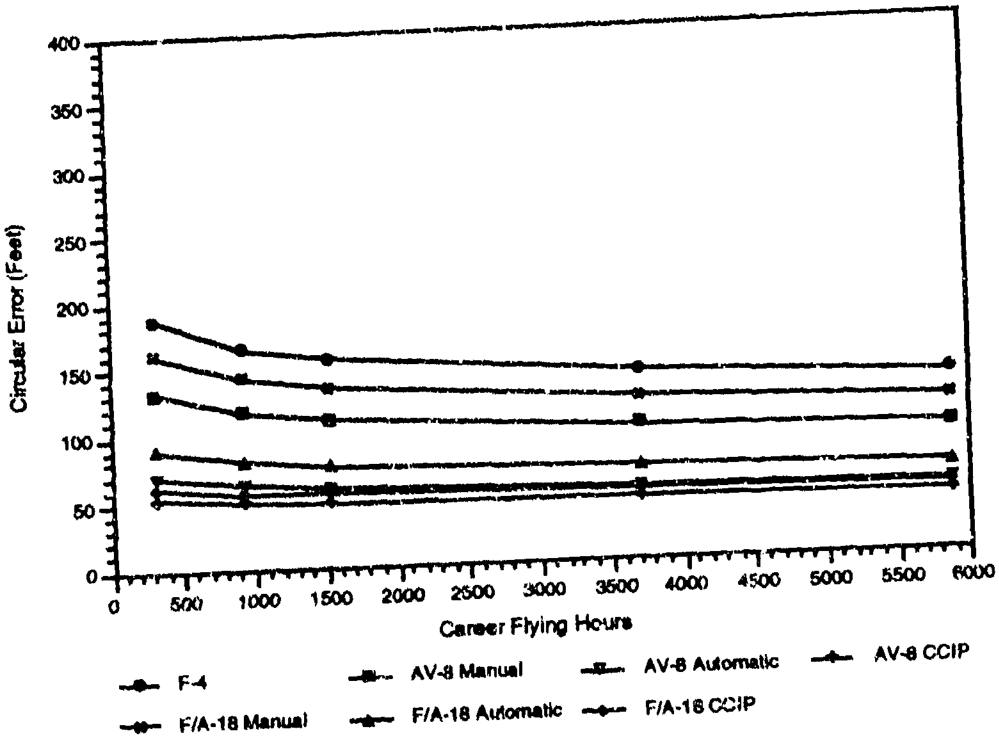

Figure 1. Bombing Error Versus Career Flying Hours (Fleet Pilots)

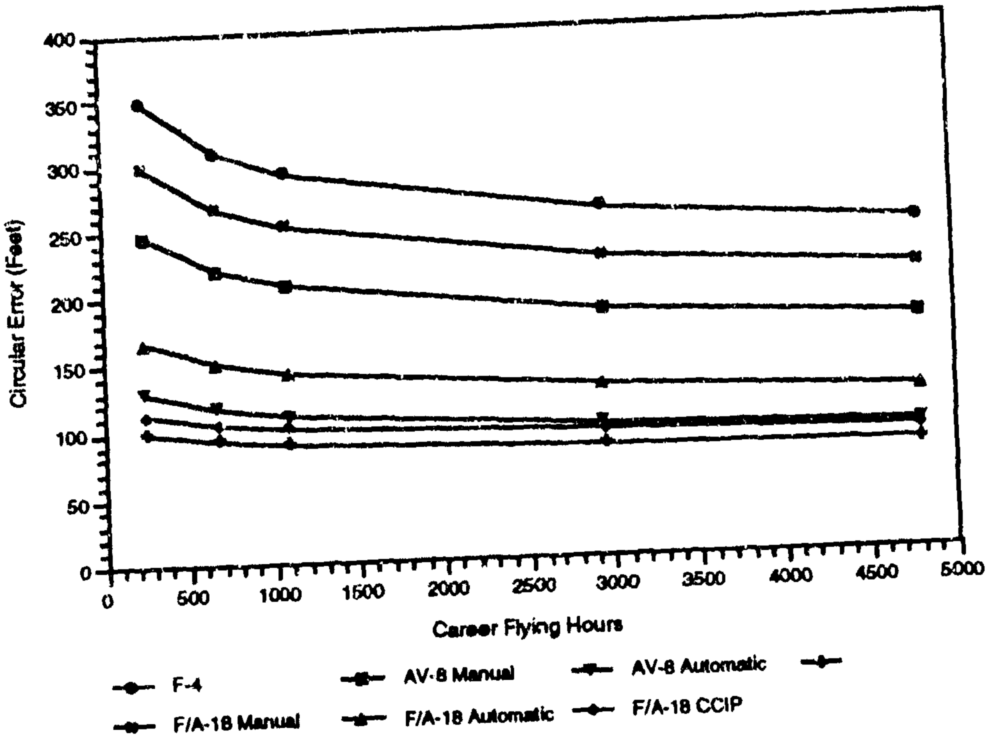

Figure 2. Bombing Error Versus Career Flying Hours (Fleet Replacement Pilots)

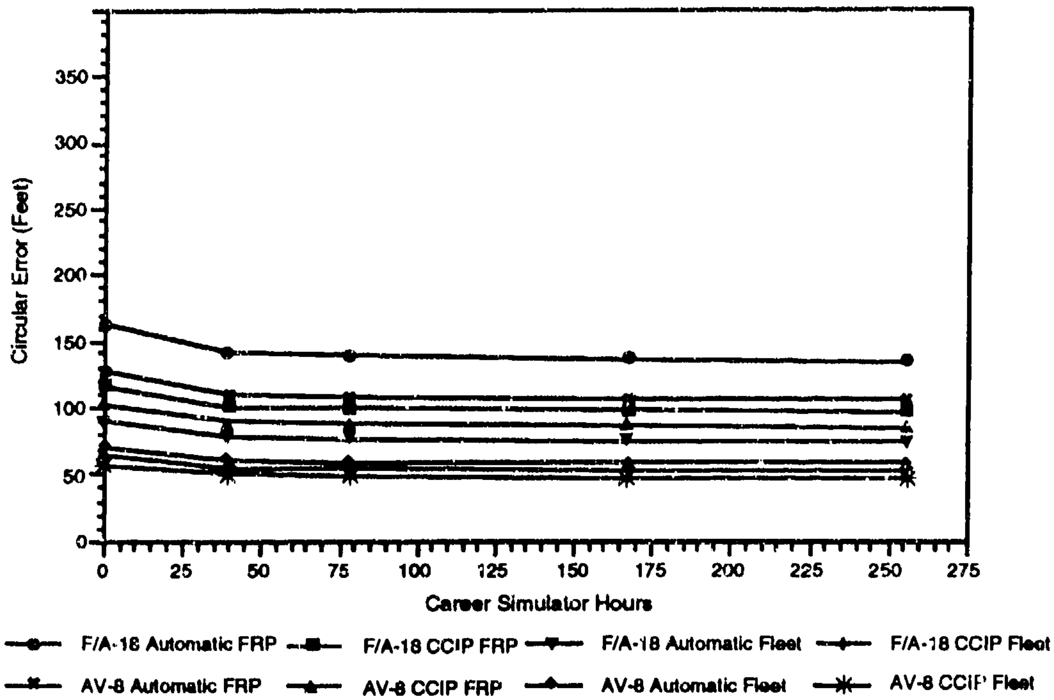

Figure 3. Bombing Error Versus Career Simulator Hours (Computed Deliveries)

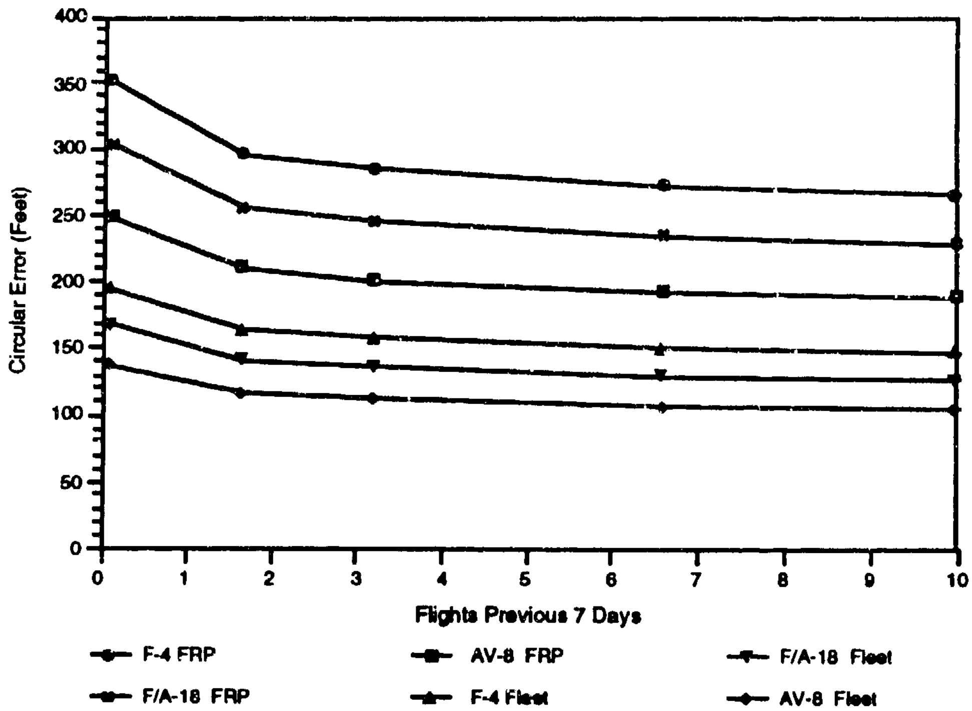

Figure 4. Bombing Error Versus Flights in the Previous 7 Days (Manual Deliveries)

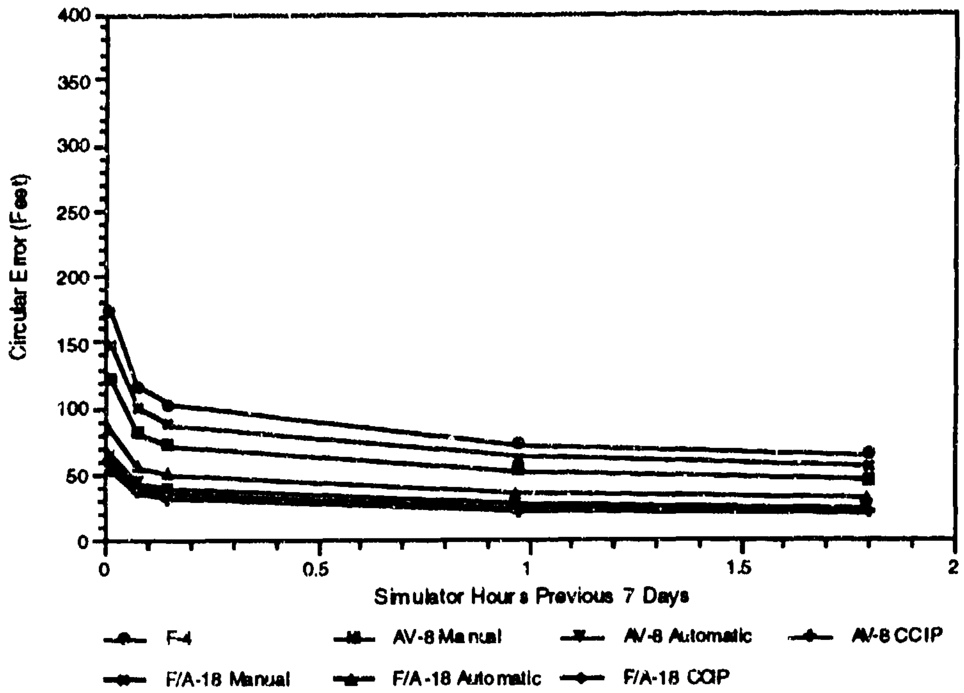

Figure 5. Bombing Error Versus Simulator Hours in the Previous 7 Days (Fleet Pilots)

The career flight-hour graphs, and the chart depicting short-term simulator hours (Figure 5) include seven curves, representing both computed and manual bombing for the three kinds of aircraft (the F-4 only performs manual bombing). Figure 3 shows eight curves—automatic and CCIP approaches for two aircraft types and both types of pilots. Figure 4 shows only six curves, two for each type of aircraft—one for fleet pilots and one for FRPs.

Two results are apparent from the graphs. Miss distances are smaller, in all cases, for fleet pilots than for FRPs for the same aircraft and approach type. The marginal improvement for an additional flying or simulator hour is greater for pilots with less experience. These are the expected results. The second result is a characteristic of the functional form of the estimated equations.

Functional forms that did not exhibit this characteristic did not fit the data as well. It has been suggested that these results indicate that pilots with more flying time should practice less, and, generally, this does guide the flying-hour program to some extent. However, the pilots with more experience are the flight leaders and instructors who train the less experienced pilots. Often they must fly to perform this job.

D. SUMMARY

The findings indicate that flying and simulator hours, both in the short-term and in the long-term, have a significant effect on bombing accuracy. These results also confirm what is generally believed about the effect on accuracy and learning of computer fire-control systems. Although the results support our intuition concerning the effect of advances in technology on learning and bombing accuracy, this does not necessarily imply that automation can replace experience. Rather, it means that less experienced pilots can master one aspect of the mission—placement of ordnance on the target—earlier in their careers. The model does not capture the effects of experience on other very important parts of the mission, such as evading opposing fire. Unless a pilot can get into and out of the target area, accurate bombing has little meaning.

The results for the short-term variables are not as robust as those for the long-term variables. The short-term effects appear selectively and depend more on the assumed functional form of the estimating equation than do the long-term effects.

According to the estimates of the coefficients of Equation (1), if the number of flights and flying hours were reduced 10% for a short period of time, the average miss distance for manual bombing runs would rise by over .5%. If the reduction continued until overall pilot experience was degraded by 10%, an additional increase of over 1% would be incurred. The long-term effect on automated delivery accuracy of a 10% decrease in flying hours would be about .75%. A similar decrease in career simulator hours would produce a decrease in accuracy of .25%. It is interesting that, at the margin, a small increase in the number of simulator hours improves bombing accuracy more than the same increase in flying hours. Since simulator hours are less expensive than flying hours, this implies that additional simulator hours are cost-effective. This, of course, may be true only for small changes near the observed mix of flying and simulator hours.

These results are largely consistent with those obtained in a 1986 Air Force study [5]. That study examined the relationships between flying experience and bombing accuracy for the A-10 and F-16 aircraft. It found both long-term and short-term experience effects with the long-term effects more pronounced than the short-term effects. They are also consistent with the results of our earlier analyses of carrier landings and air-to-air combat [4].

III. ANALYSIS OF AIR FORCE C-130 TACTICAL AIRDROPS

This section reports the results of an investigation of the relation between C-130 aircrew experience and performance in the tactical airdrop mission. This analysis is particularly interesting because it treats a multiperson mission, and because we were able to estimate the contribution of simulator hours to mission performance for one crew position. When mission performance depends on more than one person, it is generally more difficult to isolate the effect of experience on performance for each crewmember. However, for this analysis, we were able to separate the experience effects of the navigator and co-pilot.

The C-130 is a multiperson aircraft used by the Military Airlift Command (MAC) for tactical airlift. Parachute drops of personnel and equipment into designated drop zones (DZs) are an important part of the tactical mission. Proficient airdrop performance is a major tactical requirement for navigators, who are graded on every practice tactical airdrop mission. The primary objective measure of drop performance is the distance from the intended point of impact (IPI) to the parachute landing point.

The missions include parachute drops of heavy equipment, container delivery systems, personnel, and training bundles simulating these three drop types. Missions are flown by aircraft singly or in formation under both visual and instrument conditions.

The airdrop solution—aircraft course, aircraft speed, and drop time—is calculated by the navigator, based on average parachute ballistics and dead reckoning. Information used in computing the solution includes altitude, rate of fall, wind velocity and direction, temperature, and DZ characteristics. Winds are critical and wind information is derived from the pre-mission briefing, Doppler radar, observation of visual clues such as smoke, and forward spotters.

Although the navigator is considered to be the key crewmember for tactical airdrops, precise coordination is required among all members of the crew. The navigator is responsible for computing the airdrop solution, but the co-pilot cross-checks calculations and serves as primary backup. The co-pilot and navigator confirm the desired offset distance—the lateral distance between the aircraft track and the IPI. Over the DZ, the pilots are responsible for maintaining the track and offset distance from the IPI. The navigator is responsible for picking the timing points, controlling the time to release, and cross-checking the offset distance.

A. DATA

Data were obtained from Headquarters 314th Tactical Airlift Wing, MAC, Little Rock Air Force Base. The database includes mission performance data and experience data for pilots, co-pilots, and navigators from two operational squadrons flying C-130 aircraft on tactical training missions.

Airdrop performance data are maintained at the squadron level for each navigator. These data are used to evaluate and document individual qualifications and for selecting Volant Rodeo navigators, who represent their squadrons in annual MAC competitions.

Experience data are maintained at the wing level, for all aircraft types, as part of the Air Force Operations Resource Management System (AFORMS). AFORMS is an automated computer database, which includes the flying-hour history of all aircrew personnel by aircraft type and model, and individual flight records for the previous 12 months. Simulator time is included in both the individual flying-hour histories and daily flight records. Unlike the Navy/Marine NAVFLIRS database, it is not possible to reconstruct daily or even monthly flight histories back more than one year. At the end of each month, the total flight histories are updated and the individual flight entries are dropped for the earliest month (prior to one year).

Flight histories for the previous 7, 30, and 60 days were constructed from the daily flight records. Daily flight data are recorded by individual, but it is possible to reconstruct the crew for each flight using dates, flight times, and aircraft tail numbers.

Performance data were obtained for 938 low-altitude tactical mission drops.[12] This number was reduced to 808 single ship or lead aircraft mission drops. Formation drops are made in accordance with the lead aircraft navigator. Although an attempt is made to adjust the scores of following aircraft to eliminate lead navigator errors, wing staff officers do not consider these adjustments to be reliable. They suggested that these drops be eliminated from the database.

The database was further reduced to 477 observations because of missing data points. For example, for a scored drop corresponding to the earliest individual flight record in the AFORMS database, there are no short-term data (flying or simulator hours in the previous 7, 30, or 60 days). The data used in the analysis are summarized in Table 5.

---

[12] Although we could usually match the crews, we were unable to identify the co-pilot for an additional 460 observations. This prevented us from analyzing the co-pilot/navigator team effect. For example, if we were able to construct a complete history, we could include a variable that represents the number of times in a certain time period that the navigator and co-pilot had flown together.

Table 5. Data Used in the Analysis of Air Force C-130 Tactical Airdrops

Variable	Minimum	Mean	Maximum
Circular Error (yards)	0	108	450
Co-pilot Career Flying Hours	74	794	3,431
Co-pilot Career Simulator Hours	0	80	175
Co-pilot Flying Hours in the Previous 60 Day	4	61	133
Co-pilot Flying Hours in the Previous 30 Days	0	30	77
Navigator Career Flying Hours	70	1,122	3,651
Navigator Career Simulator Hours	0	72	563
Navigator Flying Hours in the Previous 60 Days	2	65	135
Navigator Flying Hours in the Previous 30 Days	0	32	80
Night Drop[a]	0	0.30	1
Heavy Equipment Drop	0	0.14	1
Training Bundle Drop	0	0.42	1
Personnel Drop	0	0.08	1
Container Delivery System Drop (Base Case)	0	0.36	1

[a] The last five variables are dichotomous. They either take the value one, showing that an observation has the indicated property, or the value zero, showing that it does not. The mean is the fraction of the observations that have the indicated property.

B. ANALYSIS

The central hypotheses of this analysis are that crewmembers with more career flying and simulator experience drop more accurately and that greater recent flying experience is associated with more accurate drops. Subsidiary hypotheses are that deliveries made at night are less accurate than drops during the day and that accuracy varies over different drop types.

1. The Tobit Censored Regression Model

The data show that drop scores, expressed as the distance in yards of the parachute landing point from the point of intended landing (circular error) are distributed continuously for drops outside of the 25-yard circle. However, the only score recorded at a range less than 25 yards was zero (bull's-eye). This implies a mixed distribution, one which is discrete for drops within 25 yards, with mass concentrated at the zero point, and continuous beyond 25 yards. From the appearance of the data, we postulate that a score of zero indicates a drop inside the 25-yard circle, but are not sure exactly where within the circle.[13] This suggests a Tobit (censored regression) model. The Tobit model is defined by Equation (2).

---

[13] This was confirmed by the wing staff. Hits outside the 25-yard circle are scored as close to the actual distance as possible. Hits within 25 yards are scored as a bull's-eye.

$$y_i^* = B'x_i + u_i \tag{2}$$
$$y_i = y_i^* \text{ if } y_i^* > C$$
$$y_i = 0 \text{ otherwise}$$

where:

y_i^* = the true values of the dependent variable (circular error in yards for this analysis)

B = a vector of unknown parameters

x_i = a vector of independent variables

u_i = residuals distributed $N(0, s^2)$

y_i = the observed (reported) values of the dependent variable

C = the censoring point (in our case 25 yards).

The censored regression model is described by Amemiya [6], Maddala [7], and Greene [8].

An examination of the raw data shows that the observations beyond 25 yards appear to be distributed as a truncated lognormal variable. Furthermore, if the observations scored as zero were spread between 0 and 25 in a lognormal pattern, the overall distribution would be lognormal. The lognormal distribution is related in a simple way to the more familiar bell-shaped normal distribution. The random variable y is said to have the lognormal distribution with mean $\exp\{m + (1/2)s^2\}$ if the natural logarithm of y (Lny) is distributed normally with mean m and variance s^2.

The Tobit procedure treats the observations as if they were normally distributed, with the censored observations falling inside the lower tail in a normal pattern rather than all bunched at a single point.[14] The Tobit model was estimated using the Davidon/Fletcher/Powell maximum-likelihood algorithm [9, p.176].

The hypotheses stated above and the analysis of the raw data led to the formulation of the following estimating equation:

$$\begin{aligned} LnCE = {} & b_0 + b_1 \times H_{cpt} + b_2 \times H_{cpst} + b_3 \times H_{cp60} + b_4 \times H_{nt} + b_5 \times H_{nst} \\ & + b_6 \times H_{n60} + b_7 \times N + b_8 \times D_{be} + b_9 \times D_{tb} + b_{10} \times D_{pers} \end{aligned} \tag{3}$$

where:

Ln = the natural log

[14] We estimate the natural log of CE, which is distributed as a normal variate.

CE = drop accuracy (circular error), the distance in yards by which the parachute misses the IPI

H_{cpt} = co-pilot career flying hours

H_{cpst} = co-pilot career flight simulator hours

H_{cp60} = co-pilot flying hours in the previous 60 days

H_{nt} = navigator career flying hours

H_{nst} = navigator career flight simulator hours

H_{n60} = navigator flying hours in the previous 60 days

N = a dummy variable taking the value 1 for a night drop and 0 otherwise

D_{he} = a dummy variable taking the value 1 for a heavy equipment drop and 0 otherwise

D_{tb} = a dummy variable taking the value 1 for a training bundle drop and 0 otherwise

D_{pers} = a dummy variable taking the value one for a personnel drop and 0 otherwise

b_0 through b_{10} are coefficients to be estimated.

The coefficients b_1 and b_4 measure the effect on drop performance of additional co-pilot and navigator career flying hours, respectively, holding career simulator hours constant. Both b_1 and b_4 are expected to be negative—better performance is reflected in smaller miss distances. The coefficients b_2 and b_5 measure the effect on mission performance of co-pilot and navigator career simulator hours, respectively, holding career flying hours constant. We conjecture that circular error will decrease if either career flying hours or simulator hours are increased while holding the other constant. We expect b_2 and b_5 to be negative. The coefficients b_3 and b_6 measure the effect on drop accuracy of co-pilot and navigator flight time in the previous 60 days, respectively. We expect b_3 and b_6 to be negative.

N, D_{he}, D_{tb}, and D_{pers} are control variables. Night missions are considered to be more difficult than day, and we expect b_7 to be positive. The coefficients of D_{he}, D_{tb} and D_{pers} are measured relative to container delivery system drops—the only drop type not included in the equation. We cannot predict the sign of the coefficient of D_{he} (b_8), based on our interviews with experienced flight personnel. However, the hypothesized signs of the coefficients of D_{tb} (b_9) and D_{pers} (b_{10}) are straightforward. The coefficient of D_{tb} is expected to be positive because the wind affects a 30-pound training bundle much more

than the heavier container delivery system loads. Since wind is one of the most critical and least predictable factors, these drops are expected to be less accurate. We expect the coefficient of D_{pers} to be negative since personnel are able to control their impact point to some extent.

2. The Logit Model

In order to gain additional insight into the effect of experience on the probability of scoring inside the 25 yard circle, we also conducted a Logit analysis. The binomial Logit model is defined by equation (4).

$$\log\{P/[1 - P]\} = B'x_i + u_i \tag{4}$$

where:

- P = the probability of success (in this case, scoring within the 25-yard circle)
- B = a vector of unknown parameters
- x_i = a vector of observations on the independent variables—the same observations of the independent variables used to estimate Equation (3).
- u_i = a vector of residuals distributed $N(0, s^2)$.

The binomial Logit model has the characteristic that the predicted probability of success is constrained to be between zero and one, and leads to the probability of success being an S-shaped function of the independent variables. The observed value of the success variable is a dummy variable, D_{25}, which equals 1 if the drop scored within the 25-yard circle and 0 otherwise. The Logit coefficients are estimated using maximum likelihood techniques. The binomial Logit is described by Amemiya [6], Maddala [7], and Greene [8].

C. RESULTS

The estimated coefficients of Equations (3) and (4) are reported in Table 6.[15] The two co-pilot experience variables, H_{cpt} and H_{cpst}, cannot be estimated together because they are co-linear. This occurs because co-pilots with more flying hours tend to have more simulator hours as well. As a result, we were unable to separate the individual effect of each variable. The equations reported in Table 6 include only the co-pilot flying-hour

---

[15] Note that coefficients of the Tobit and Logit equations are opposite in sign. For the Tobit equation, a negative sign indicates better performance—the effect is to decrease miss distance. For the Logit equation, a positive sign means better performance—the effect is to increase the probability of landing within the 25-yard circle.

variable. This means that the estimated effect of more flying hours includes the effect of additional simulator hours that usually go along with them. A model that does unscramble these two effects is discussed in subsection C.

Table 6. Determinants of C-130 Drop Accuracy for Lead Aircraft (Tobit and Logit Estimates)

	Dependent Variable:	
Independent Variable:	Ln(CE) Tobit Model	D_{25} Logit Model
Constant	4.51 (32.1)***	-3.27 (5.89)***
Co-pilot Career Flying Hours (H_{cpt})	-0.10924E-03 (1.79)*	0.33198E-03 (1.49)
Navigator Flying Hours in the Previous 60 Days (H_{n60})	-0.33751E-02 (2.22)**	0.20110E-01 (3.27)***
Night Flight (N)	0.25005 (2.98)***	-0.59405 (1.70)*
Heavy Equipment Drop (D_{he})	0.48735 (4.29)***	-1.06939 (1.91)*
Training Bundle Drop (D_{tb})	0.27919 (3.28)***	-0.57125 (1.82)*
Personnel Drop (D_{pers})	-1.0149 (8.48)***	2.6134 (6.56)***
Standard Deviation (s)	0.7925 (24.6)***	
Number of Observations	477	477
Partial (CE or D_{25} wrt H_{cpt})[a]	-0.0134	0.435E-04
Partial (CE or D_{25} wrt H_{n60})[a]	-0.3657	0.264E-02

Note: Numbers in parentheses are t-statistics.
[a] Calculated at the mean.
* Significant at the .10 level.
** Significant at the .05 level.
*** Significant at the .01 level.

The partial effects of circular error—and of the probability of scoring within the 25-yard circle—with respect to the experience variables are also shown in Table 6. The partial effect is the marginal change in CE for the Tobit equation—the marginal change in the Prob[CE < 25] for the Logit equation—associated with a one-hour change in either of the flying-hour variables.

1. Tobit Results

Neither the short-term co-pilot nor the long-term navigator variables were significantly related to performance. The coefficients of both the long-term co-pilot and

short-term navigator flying-hour variables are statistically significant in the expected direction.[16] The coefficients of the control variables are statistically significant with the expected sign where we were able to predict the sign. We have no hypothesis concerning heavy equipment drop accuracy compared to container delivery system drops except that we expect them to be different. As it turns out, the coefficient of D_{he} is positive and highly significant.

2. Logit Results

Equation (4) was estimated using the Newton algorithm [9, p. 151]. The coefficient of the navigator short-term flying-hour variable is statistically significant in the expected direction. The coefficient of co-pilot total flying hours has the expected sign, but is only significant at the .12 level. The partial effects of miss distance with respect to all experience variables have the expected signs. The coefficients of the control variables are statistically significant at the .1 level or better and have the expected signs.

The estimated Tobit results shown in Table 6 are presented graphically in Figures 6 and 7. Figure 6 depicts the effect of changing co-pilot career flying hours, and figure 7 shows the effect of changing navigator flying hours in the previous 60 days.[17]

---

16 We estimated the equation using different experience variables for the co-pilot and navigator. The variables shown in Table 6 reflect those combinations of variables that influence accuracy and do not exhibit serious colinearity. The short-term experience variable was flying hours or flights in the previous 7 days for the analysis of bombing accuracy. The 60-day variable was more successful in Equation (3). There is no intrinsic reason why recently honed skills ought to depreciate at the same rate for all flying activities. It is, however, interesting that different aircraft, crew positions, and mission equipment are sensitive to differences in the choice of a short-term experience variable. The issue of skill depreciation for different aircraft and missions should perhaps be pursued in follow-on research.

17 In order to graph the estimated equation we made the transformation:

$CE = \exp\{LnCE\} = \exp\{\Sigma_i b_i x_i\}$.

However, this is a biased estimate of CE because the mean of the transformed residual, which is non-zero, is omitted. Recall that for Lnu_i distributed $N(0,s^2)$, then u_i is distributed lognormal with mean $\exp\{0 + (s^2/2\}$. We therefore include the correction $\exp\{s^2/2\}$ where s is the estimated standard deviation calculated by the Davidon/Fletcher/Powell algorithm used to estimate the Tobit model. The transformed equation is therefore:

$CE = \exp\{\Sigma_i b_i x_i\} \exp\{s^2/2\}$.

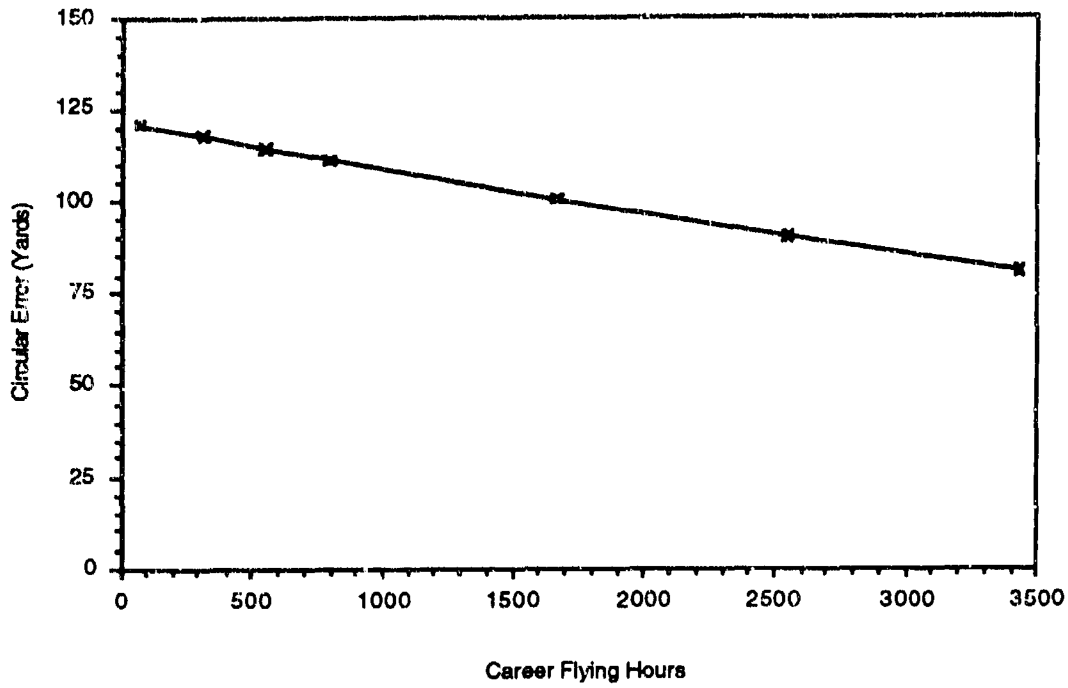

Figure 6. C-130 Tactical Drop Error Versus Co-pilot Career Flying Hours

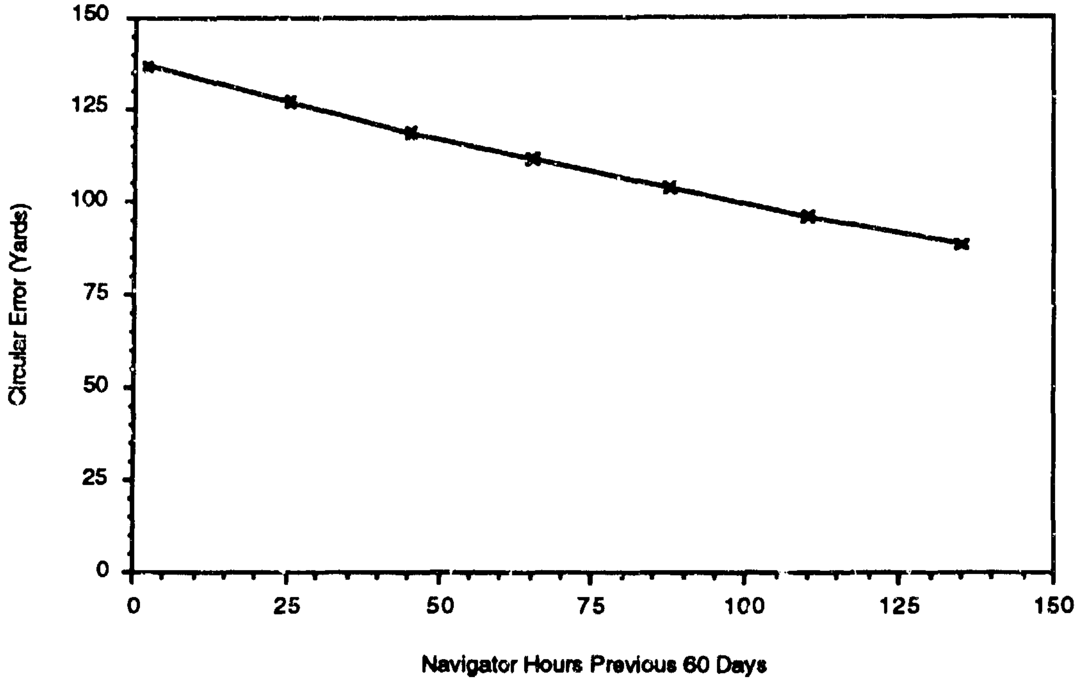

Figure 7. C-130 Tactical Drop Error Versus Navigator Flying Hours in the Previous 60 Days

D. CO-PILOT SIMULATOR HOURS: AN EXCURSION

In order to unscramble the individual effects of the co-pilot flying and simulator hours, the ratio of co-pilot simulator to flying hours was added to the right side of Equation (3). The interpretation of this variable is straightforward. We estimate the effect of changing flying hours while holding the ratio of simulator hours to flying hours constant. We also estimate the effect of changing simulator hours while holding flying hours constant. Although all other terms were statistically significant with the expected signs, the ratio term had the correct sign but was not statistically significant. We then estimated the following equation.

$$\begin{aligned} LnCE = \; & b_0 + b_1 \times H_{cpt} + b_2 \times (LnH_{cpst}/LnH_{cpt}) + b_3 \times H_{n60} + b_4 \times N \\ & + b_5 \times D_{he} + b_6 \times D_{tb} + b_7 \times D_{pers} \end{aligned} \quad (5)$$

The variable definitions for Equation (5) are the same as for equation (3) and b_0 through b_7 are the coefficients to be estimated.

The hypothesized signs of the coefficients are the same as for Equation (3), with the addition of a conjectured negative sign for the coefficient (b_2) of the ratio term. In this model, the coefficient of H_{cpt} measures the effect on miss distance of changing co-pilot career flying hours while holding the ratio of the natural logs of simulator and flying hours constant. The numeric value of the coefficient b_2 is not as easy to interpret as it would be for the ratio H_{cpst}/H_{cpt}. However, since the natural log is a monotonically increasing function of the variable itself, the two ratios move in the same direction with changes in H_{cpst} or H_{cpt}. The coefficient b_2 then measures the effect on miss distance of a change in the natural log of simulator hours while holding co-pilot career hours constant. We can therefore measure the change in miss distance associated with a change in either flying or simulator hours.[18]

The estimated Tobit and Logit coefficients of Equation (5) are reported in Table 7.

---

[18] The ratio term LnH_{cpst}/LnH_{cpt} is correlated much less with H_{cpt} than is H_{cpst} itself. Although the significance of the ratio term coefficient indicates that conditioning of the data matrix is not a problem, we tested the model for ill-conditioning of the data matrix (X), using the singular-value decomposition of X method described by Belsley, Kuh, and Welsch [10]. The results of this test—all condition numbers less than 8 and no large correlations—indicate that the X matrix is well-conditioned. The data matrix using this formulation is also better conditioned than for the equation using the ratio of simulator to flying hours.

Table 7. Determinants of C-130 Drop Accuracy for Lead Aircraft: An Excursion (Tobit and Logit Estimates)

Independent Variable:	Dependent Variable: Ln(CE) Tobit Model	Dependent Variable: D_{25} Logit Model
Constant	4.99 (15.6)***	-6.66 (4.89)***
Co-pilot Career Flying Hours (H_{cpt})	-0.16113E-03 (2.31)**	0.74676E-03 (2.83)***
Ratio of the Logs of Co-pilot Simulator to Actual Flying Hours ($Ln(H_{cpst})/Ln(H_{cpt})$)	-0.64142 (1.69)*	4.5110 (2.77)***
Navigator Flying Hours in the Previous 60 Days (H_{n60})	-0.35265E-02 (2.35)**	0.19507E-01 (3.12)***
Night Flight (N)	0.23297 (2.79)***	-0.47437 (1.35)
Heavy Equipment Drop (D_{he})	0.49115 (4.38)***	-1.1303 (2.00)**
Training Bundle Drop (D_{tb})	0.29107 (3.40)***	-0.74619 (2.26)**
Personnel Drop (D_{pers})	-0.97429 (8.04)***	2.3998 (5.93)***
Standard Deviation (s)	0.7891 (24.6)***	
Number of Observations	477	477
Partial (CE or D_{25} wrt H_{cpt})	-0.89E-02	0.247E-04
Partial (CE or D_{25} wrt H_{cpst})	-0.1311	0.111E-02
Partial (CE or D_{25} wrt H_{n60})	-0.3851	0.256E-02

Note: Numbers in parentheses are t-statistics.
* Significant at the .10 level.
** Significant at the .05 level.
*** Significant at the .01 level.

1. Tobit Results

The coefficients for the long-run co-pilot and short-run navigator flying-hour variables are statistically significant in the expected direction, as is the coefficient of the ratio of the natural logs of co-pilot simulator and flying-hour term. The coefficients of the control variables are statistically significant with the expected signs.

The partial effect of miss distance with respect to co-pilot simulator hours is -.1311 compared to -.0089 for co-pilot flying hours. This implies that an additional simulator hour reduces miss distance by more than an additional flying hour. Although this implies a substitution effect between flying and simulator hours, it does not mean that the apparent greater benefit of career simulator hours will hold true except near the observed

levels of the independent variables. As Table 5 shows, co-pilots average ten times more flying hours than simulator hours.

2. Logit Results

The coefficients of all experience variables, including total co-pilot flying hours, are significant at the .01 level and have the expected signs. The coefficients of the control variables are statistically significant at the .05 level except for the night dummy, and all have the expected signs.

The model specified by Equation (5) allows us to separate the effects of co-pilot flying and simulator hours, and therefore to gain some insight into the substitution effect between flying and simulator hours. However, the functional form has one undesirable mathematical property. The miss distance is not a monotonically decreasing function of co-pilot career flying hours throughout the range of co-pilot flying hours. This means that the estimated equation predicts that miss distance increases with increasing co-pilot flying hours for values of flying hours less than approximately 300. For larger values, which includes most of the observations, the function is well behaved, and miss distance decreases monotonically with increasing co-pilot flying hours.[19]

The Tobit results for co-pilot simulator hours are shown graphically in Figure 8.

E. SUMMARY

Our analysis of C-130 tactical airdrop performance shows that, as with Marine bombing accuracy, flying time—both short- and long-term—makes a significant difference in performance.

There is an important difference in these results, compared with the analysis of Marine Corps bombing, as well as with other programs we have analyzed. In those analyses, most of the impact of changes in experience on performance appeared to operate through the long-term experience variables. Although the long-term variable is more important than the short-term for the co-pilot—flying hours in the previous 30 or 60 days was not a significant variable—this result was reversed for the navigator.

---

[19] Miss distance does decrease monotonically with increasing co-pilot career simulator hours throughout the range of both flying and simulator hours. In addition, if the ratio of simulator hours to flying hours is held constant, as is the case if we assume the current simulator/flying-hour mix, the function is well-behaved with respect to changes in co-pilot career flying hours.

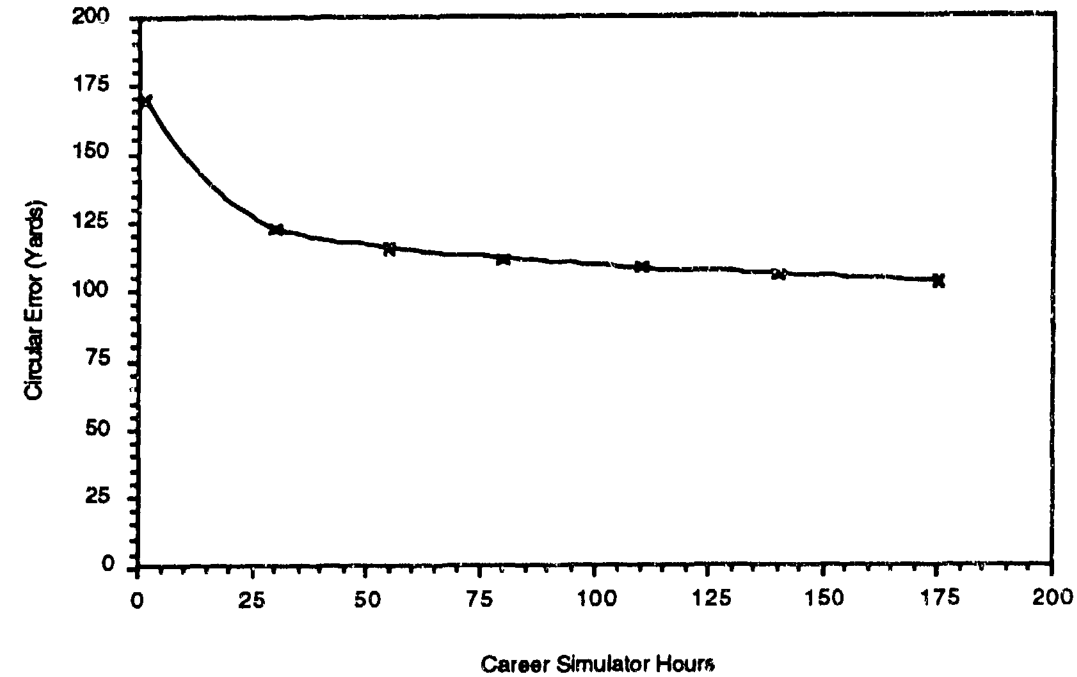

Figure 8. C-130 Tactical Drop Error Versus Co-pilot Career Simulator Hours

For the C-130 analysis, the short-term navigator variable is a more important determinant of performance than any of the long-term variables we examined. In fact, the long-term navigator variable is not a significant determinant of drop accuracy. Although this is different from our other findings, it is in agreement with the intuition of experienced navigators we interviewed. Based on their experience, they expected the short-term navigator variable to have a much greater effect on performance than career flying experience for either the navigator or the co-pilot. The primary reason given for this was that tactical air drops, though an important part of the mission, are only one aspect of the overall navigator mission. The navigators offered the explanation that if we could examine missions that more nearly reflect the total demand on navigators, the effect of long-term experience would show up. One factor mentioned was that many of our observations represent flights to familiar ranges. The navigators are therefore not fully tested in their ability to navigate into the DZ and execute the mission in unfamiliar wind and terrain conditions.

A second important difference between these findings and those of Reference [4] is that co-pilot simulator hours have a significant effect on performance. As Table 7 shows, the marginal (partial) effect at the mean of the data is estimated to be greater for simulator hours than for flying hours. Since simulator hours are less expensive than flying hours,

this implies that at the current mix and levels of flying and simulator hours, additional simulator hours are cost-effective. Keep in mind that these results are only valid in the range of the observations.

Figures 9 and 10 show the effect on miss distance, in three dimensions, of changes in co-pilot career flying and simulator hours. Figure 9 shows the effect of a 10-hour change in flight and/or simulator time centered on the variable means. Figure 10 is an expanded view. These graphs show the potential for changing average miss distance by changing co-pilot career flight or simulator hours individually or in combination.

The model defined by Equation (5) provides us with both an analytic and a policy tool. It enables us to unscramble the effects of simulator and flight hours. It also gives us a means of analyzing the relative cost-effectiveness of each.[17]

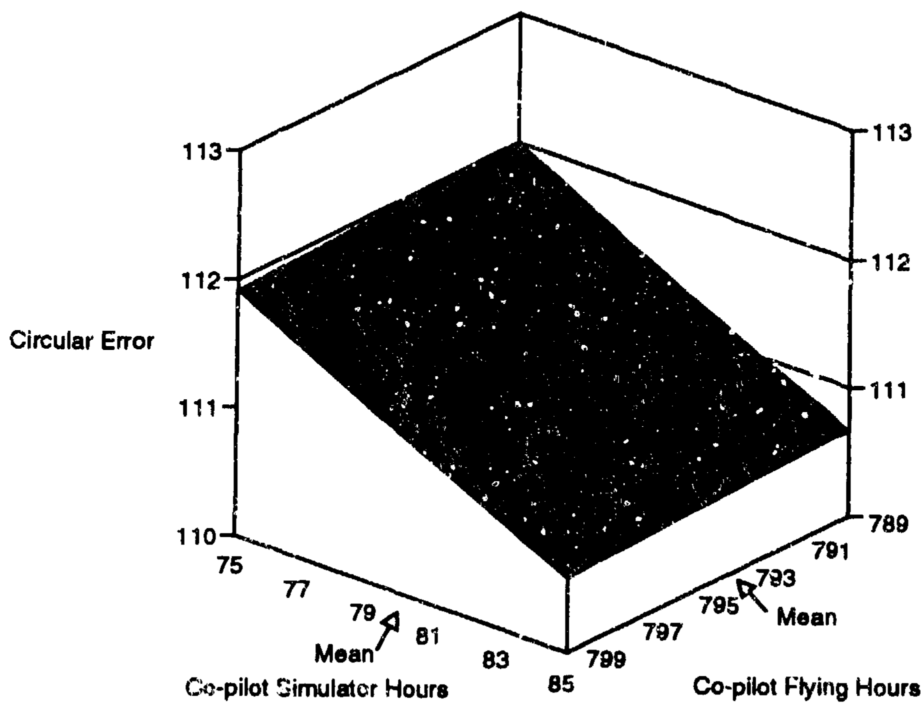

Figure 9. Change in Miss Distance with Changes in Co-pilot Flying and Simulator Hours at the Variable Means

---

[20] It is speculative whether the representation shown in Figure 10 would hold true throughout the full range of both variables without further analysis. Because of the existing limited variation in the ratio of simulator to flying hours, it would probably be necessary to conduct some controlled experimentation to fully test the effects shown in Figure 10.

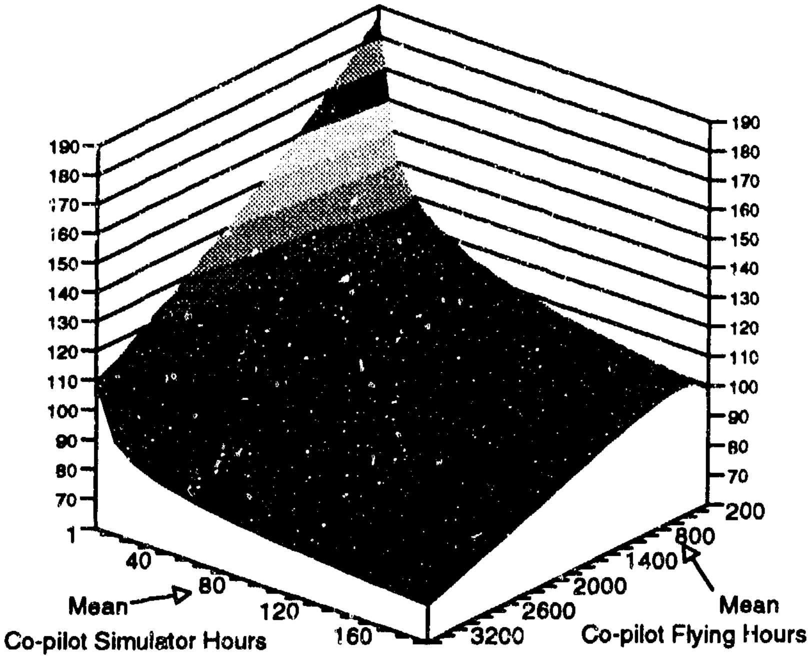

Figure 10. Change in Miss Distance with Changes in Co-pilot Flying and Simulator Hours (Expanded View)

IV. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

Existing data on performance can be used to develop relationships between aircrew performance and both long-term and short-term experience variables that reflect the impact of variations in the flying-hour program. The following conclusions can be drawn from the analyses presented here:

- The accuracy of Marine Corps bombing is influenced by the number of hours flown by pilots, both recently, and over the course of their careers. Pilot flight simulator hours influence performance significantly for Marine Corps bombing, and simulator hours appear to be cost-effective in this instance.
- A 10% reduction in flying hours, both career and recent, is estimated to decrease bombing accuracy by about 2% for manual bomb deliveries. For computed bomb deliveries, a 10% decrease in flying hours is estimated to increase the average circular error by only .75%.
- The performance of MAC aircrews in tactical air drops is significantly related to co-pilot career flying and simulator hours, and simulator hours appear to be cost-effective in this instance. We were not able to show a significant effect of simulator hours on performance for other crew positions. The number of hours flown by the navigator in the previous 60 days is a significant determinant of crew performance. The apparent absence of an effect on performance of navigator career hours is an anomaly. This may be explained by the fact that the observed drops do not fully capture the navigator's contribution to the total mission.
- If C-130 navigator flying hours were reduced 10% for a period of 2 months, the average tactical airdrop miss distance would increase by approximately 2.5%. If the reduction in total flying hours continued until total co-pilot experience was degraded by 10%, holding co-pilot simulator hours constant, we would expect an additional increase in miss distance of approximately .5%. If co-pilot simulator hours were decreased by 10%, holding flying hours constant, the expected increase in circular error would be nearly 1%. The elasticity, at the mean, of miss distance with respect to simulator hours is greater than for flying hours by a factor of two, but the marginal effect is much greater. This means that, at the current mix of flying and simulator hours, we

estimate a substantially greater return for an additional simulator hour than for an additional flying hour.

- If C-130 navigator flying hours were reduced, in the short-term, by 10%, the probability of a tactical airdrop landing within a 25-yard circle would decrease by approximately 11%. If the reduction in total flying hours continued until co-pilot total experience was degraded by 10%, holding co-pilot simulator hours constant, an additional decrease of 1% would be incurred. The effect of a 10% decrease in co-pilot simulator hours, holding flying hours constant, would be to decrease this probability by 6.5%.
- For Marine bombing, the long-term effects of flying hours on performance appear to be quantitatively more important than the short-term effects. This is consistent with previous findings for Navy carrier landings and air-to-air combat [4]. This means that in an emergency it would be difficult to remedy an inadequate level of training. The availability of aircraft and of training ranges would also constrain the ability to improve performance quickly. This is in agreement with the findings of our previous analyses and for the C-130 co-pilot flying hours. It is not in agreement with our findings for C-130 navigators.

Table 8 summarizes the impact on performance of a 10% reduction in flying hours from the current level. The carrier landing and air-to-air combat findings from Reference [4] are repeated for comparison.

Table 8. Impact on Performance of 10% Cuts in Flying-Hour Variables

Performance Measure	Career Flying	Recent Flying	Total Flying	Career Simulator
Marine Corps Bombing Miss Distance	1.2%	0.6%	1.8%	0.25%
C-130 Tactical Airdrops				
Co-pilot	0.5%		0.5%	1.0%
Navigator		2.3%	2.3%	
Total	0.5%	2.3%	2.8%	
Unsatisfactory Landings[a]	6.9%	2.6%	9.5%	
Air-to-Air Combat[a,b]				
Probability Red Kills Blue	6.3%	2.9%	9.2%	
Probability Blue Kills Red	-2.6%	-2.2%	-4.8%	

[a] Source: Reference [4].
[b] Red are aggressor aircraft; Blue are friendly aircraft.

Existing evidence can be used to help justify flying-hour programs, since it shows that flying hours make an objectively measurable difference, but the results are not easily generalized. Results differ considerably by mission and aircraft, and many missions and

aircraft have not been analyzed. Also, the wisdom of spending more on flying hours depends not only on the improvement in performance, but on both the value of that improvement and the cost of achieving it in different ways.

B. RECOMMENDATIONS

A start has been made in linking flying hours to indicators of aircrew proficiency. Analysis of objective data has shown that, in all the cases studied, reductions in flying hours will lead to measurable degradations in mission performance. Still, there is much work to be done. The kinds of relationships developed in this paper are not, by themselves, adequate for the task of determining how many flying hours are enough.

Further studies along the lines of this analysis are needed to determine quantitative relationships between short- and long-term flying and simulator hours and performance for a broader range of missions, aircraft types, and crew positions. Such analyses should include:

- Studies designed to gain a better understanding of the cost-effectiveness implications of changes in flying-hour budgets for different missions, aircraft types, and crew positions.
- Analyses aimed at further documenting the combined effects of simulator and actual flying hours on performance.

Little analysis has been performed that addresses the cost-effectiveness of flying-hour budget reductions. Savings in flying-hour budgets may result in higher costs in other areas. These higher costs may be borne in peacetime or may not show up except in combat. The most obvious additional peacetime cost is that associated with increased accident rates. Planes and pilots lost in accidents must be replaced, and both aircraft and pilot training are very expensive. In addition, a recurring observation made by aircrew personnel during our interviews was that they must fly with some minimum frequency, which they perceive as a safe level, or they will not remain in the service. Reduced aircrew retention increases training costs. A reduced level of aircrew experience also adversely affects readiness.

Decreased performance in combat is likely to be associated with lost aircrew personnel, aircraft, and ships. Achievement of combat objectives may require larger forces or more sorties. These costs should be measured and compared with proposed flying-hour budget savings. Both Marine Corps bombing and air-to-air combat lend themselves to this type of analysis.

The second type of analyses should address the interaction between flying hours and simulator hours.

Our work has tended to support the importance of adequate flying-hour budgets, but the defense budget is extremely tight today. Our preliminary findings on the cost-effectiveness of simulator time needs to be examined further. While statistical studies of additional aircraft are desirable, our research indicates that lack of variability in the pattern of simulator use may make analysis difficult. A series of experiments that varies simulator use across crewmembers is needed to clarify the proper mix of flying time and simulator time.

REFERENCES

[1] General Accounting Office. "The Defense Budget: A Look at Budgetary Resources, Accomplishments and Problems." GAO/PLRD-83-62, April 1983.

[2] General Accounting Office. "Aircrew Training: Tactical Air Command and Strategic Air Command Flying Hour Programs." Briefing Report to the Chairman, Subcommittee on Defense, Committee on Appropriations, House of Representatives, United States General Accounting Office, September 1986.

[3] Horowitz, Stanley A., Colin P. Hammon, and Paul R. Palmer. "Relating Flying-Hour Activity to the Performance of Aircrews: A Progress Report." Institute for Defense Analyses, Paper P-2085, December 1987.

[4] Hammon, Colin P., and Stanley A. Horowitz. "Flying Hours and Aircrew Performance." Institute for Defense Analyses, Paper P-2379, October 1990.

[5] Cedel, Lt. Col. Thomas E., and Lt. Col. Ronald P. Fuchs. "An Analysis of Factors Affecting Pilot Proficiency." Fighter Division, Directorate of Theater Force Analyses, Air Force Center of Studies and Analyses, December 1986.

[6] Amemiya, Takeshi. "Qualitative Response Models: A Survey," *Journal of Economic Literature*, Vol. 19 (December 1981), pp. 1483-1536.

[7] Maddala, G. S. *Limited Dependent and Qualitative Variables in Econometrics.* Cambridge, MA: Cambridge University Press, 1983.

[8] Greene, William H. *Econometric Analysis.* New York, NY: Macmillan Publishing Co., 1990.

[9] Greene, William H. "LIMDEP." Department of Economics, Graduate School of Business Administration, New York University, 1985.

[10] Belsley, David A., Edwin Kuh, and Roy E. Welsch. *Regression Diagnostics: Identifying Influential Data and Sources of Colinearity.* New York, NY: John Wiley & Sons, 1980.

ABBREVIATIONS

AFORMS	Air Force Operations Resource Management System
CCIP	Continuously Computed Impact Point
CEP	circular error probable
DZ	drop zone
FRP	Fleet Replacement Pilot
HUD	head-up display
IDA	Institute for Defense Analyses
IPI	intended point of impact
MAC	Military Airlift Command
NAVFLIRS	Naval Flight Record Subsystem